European Journal of Human Genetics (1998) 6, 417–429

http://www.stockton-press.co.uk/ejhg

ORIGINAL PAPER

Parallel molecular genetic analysis

Steven E McKenzie[1,2], Elaine Mansfield[3], Eric Rappaport[4], Saul Surrey[1] and Paolo Fortina[4]

[1]*Department of Pediatrics, Thomas Jefferson University, Philadelphia, PA and the du Pont Hospital for Children, Wilmington, DE*
[2]*Department of Chemical Engineering, University of Pennsylvania School of Engineering and Applied Science*
[3]*DiaDexus, 3303 Octavius Drive, Sante Clara, CA, USA*
[4]*Department of Pediatrics, University of Pennsylvania School of Medicine, Children's Hospital of Philadelphia, Philadelphia, PA, USA*

We describe recent progress in parallel molecular genetic analyses using DNA microarrays, gel-based systems, and capillary electrophoresis and utilization of these approaches in a variety of molecular biology assays. These applications include use of polymorphic markers for mapping of genes and disease-associated loci and carrier detection for genetic diseases. Application of these technologies in molecular diagnostics as well as fluorescent technologies in DNA analysis using immobilized oligonucleotide arrays on silicon or glass microchips are discussed. The array-based assays include sequencing by hybridization, cDNA expression profiling, comparative genome hybridization and genetic linkage analysis. Developments in non microarray-based, parallel analyses of mutations and gene expression profiles are reviewed. The promise of and recent progress in capillary array electrophoresis for parallel DNA sequence analysis and genotyping is summarized. Finally, a framework for decision making in selecting available technology options for specific molecular genetic analyses is presented.

Keywords: chip DNA; microarray; mRNA level; capillary array electrophoresis; single nucleotide polymorphism; mutation detection; genetic analysis

Introduction

Progress from the Human Genome Initiative and related efforts will soon result in the generation of the sequence of all expressed human genes and the identity of many more disease-causing or disease-predisposing mutations. Knowledge of genomes of multiple other eukaryotic and prokaryotic organisms has progressed as well. Increasing attention has been devoted to parallel molecular genetic analysis methods in order to facilitate the search for multiple genetic alterations at the same time and to advance functional genomics (assessment of gene expression in specific tissues, during development, or in disease states). This review summarizes recent developments in parallel molecular genetic analysis of interest to human genetics.

Functional genomics studies increasingly necessitate obtaining the results of multiple molecular genetic analyses at the same time. For example, genes which

Correspondence: Paolo Fortina, MD, The Children's Hospital of Philadelphia, 310-C Abramson Pediatric Research Center, 34th Street and Civic Center Boulevard, Philadelphia, PA 19104-1804, USA. Tel: (215) 590 3318; Fax: (215) 590-3660; Email: fortina@mail.med.upenn.edu
Received 10 November 1997; revised 2 March 1998; accepted 31 March 1998

cause human diseases may contain many different mutations linked to the phenotype, and efficient strategies to scan for these mutations are needed. Furthermore, thousands of genes are expressed in any given cell. Differences in gene expression between cells accompany both normal development and development of the malignant cell phenotype. Methods for rapid identification of differentially expressed genes will lead to increased understanding of cell development and perhaps better therapy for cancer. Parallel molecular genetic analysis provides simultaneous information about many genes. Parallel analyses can also provide large amounts of information from small amounts of starting cellular material. The concept of parallel is distinct from but related to that of high throughput, which results in the generation of large amounts of molecular genetic information per unit time. It is possible to have a parallel process which is not high throughput if it takes a long time to obtain information about multiple genes. Conversely, most high-throughput processes are parallel, because serial molecular genetic analyses will tend to increase the overall time needed. Methods in which multiple wet biochemical steps within a single tube or microwell for each molecular analysis are followed by gel electrophoresis of radiolabeled products and then development of a gel image by autoradiography have been standard in molecular genetic analysis, but such methods are intrinsically low throughput. In this review we examine several methods of parallel molecular genetic analysis which are of increasing importance in a number of fields.[1–5] DNA microarrays, capillary array electrophoresis, mass spectrometry, homogeneous solution assays and hybrids of these methods are all particularly promising technologies.

We begin with DNA microarrays with their myriad uses, then move to parallel mutation detection methods and gene expression profiling methods that do not use microarrays, and finally capillary array electrophoresis approaches to DNA sequence analysis and microsatellite-based genotype analysis. Generalizable high-throughput analytic approaches accessible to any laboratory with molecular genetics expertise will be highlighted. Areas which will not be discussed and for which the reader is directed to recent reviews or references include gene expression profiling *in situ*,[6] protein expression profiling using 2D gels and mass spectrometry analysis,[7–9] parallel cellular and nucleic acid sample preparation,[10–12] and bio-informatics approaches to data processing and interpretation.[13–15]

DNA Microarrays

DNA microarrays have been variously called DNA chips and DNA biochips. The common denominator is a solid platform with a set of immobilized nucleic acid species that participates in a solid-fluid interfacial interaction with a solution of complementary nucleic acid targets. Uses of the DNA microarray include mutation and polymorphism detection, definition of gene expression profiles, genotyping, defining gene organization and mapping, and DNA sequence analysis of previously uncharacterized regions, among others.[16–30] Since a high density of genes can be studied in parallel, only a small amount of sample is needed. This technology is mostly robotic driven; consequently it can easily be introduced into many research and clinical laboratories.

Mutation and Polymorphism Detection

Microchips consisting of ordered arrays of oligonucleotide probes have been applied to hybridization-based mutation detection schemes.[16–39] This approach represents in some sense a miniaturization of the use of immobilized oligonucleotides in 96 well plates with several important technical and conceptual differences. With glass or silicon chips, a high degree of parallel array formation increases the information content. With a high number of closely related oligos on a surface, information about partial matches may also be obtained when perfect match duplex values are obtained with fluorescence. Furthermore, the PCR step which is often used to generate reagents for annealing to immobilized probes can now be done in a silicon chip microenvironment.[40–42]

The DNA microarray technique has been used to analyze the entire human mitochondrial DNA,[24] to detect polymorphisms in the HIV-1 clade B protease gene,[25] and to detect mutations in BRCA1,[26,36] the cystic fibrosis transmembrane receptor gene and p53.[23,34] In an extension of DNA microarray technology, Pastinen and colleagues reported allele-specific detection of 12 common disease-causing mutations in the Finnish population using a microarray and single-nucleotide extension.[43] Detection of beta-globin gene alleles and mutations has also been reported by several groups.[28,44,45]

Gene Expression Profiling

The potential for this type of gene expression chip is enormous. There are at least three things one wants to know about an mRNA in a cell. Is it present or absent? What is its quantitative level of abundance? What

difference is there in this level between two different cellular samples? Parallel mRNA expression monitoring by DNA microarray methods has been reported by several different groups with different approaches. Two papers in the peer-reviewed literature summarize the Affymetrix experience in the gene expression area.[46,47] The work demonstrates detection of transcripts at the 1 in 300 000 level, equivalent to 1 copy per cell, with a dynamic range between 1000 and 10 000. Fragmented RNA target preparation appears advantageous kinetically, in that information was available with hybridization times as short as 2 hours. For the less abundant transcripts, hybridization up to 20 hours was employed. There is a very high degree of probe redundancy (20 to 300 probe pairs per target) and high complexity in the computational analysis, both of which may relate to the fact that light-directed *in situ* synthesis currently yields only 3 to 5% of full-length, correct probe sequence per spot for a 20 mer. They have not reported explicit comparison with alternative probes, such as cDNA fragments or presynthesized oligonucleotides. The generalizability of the method to any laboratory with molecular genetics expertise is limited by the need for photolithography, and adaptability for newly discovered targets is potentially slow and costly since new masks might need to be made.

The Stanford/Synteni groups have published several papers using cDNA arrays to assess gene expression, first in yeast and plants and then in human cells.[48–51] The two-color approach directly provides information comparing two cellular inputs. The cDNA microarrays are very specific and their array methodologies are adaptable by others [www.cmgm.stanford.edu/pbrown]. Alterations in a number of interesting genes in cellular pathways in response to extracellular perturbations have been identified. With cDNA probes, the kinetics of hybridization is relatively slow and the detection of low copy number transcripts, which account for the majority of the transcripts in a cell by number, is unclear. The reports cite 1 in 10 000 detected on a per mass basis, and one report indicates a 1 in 100 000 detection sensitivity on a per mass basis. With a cDNA microarray, splice variants and closely related gene family members may not be distinguished.

Workers at Molecular Dynamics reported gene expression results with their arrayer and scanner at the Hilton Head '97 meeting [www.mdyn.com/posters/hiltonhead97]. In bacterial systems, 2 to 3-fold differences in expression of transcripts at abundance levels as low as 1:300 000 were reliably detected. Custom arraying by the end user, after an initial investment, may meet the needs of many research laboratories. The Molecular Dynamics scanner employs confocal microscopy and a photomultiplier tube and is sensitive to 0.1 attomole of fluorescent dye per square micron.[52]

For microarray methods for mRNA detection and quantitation, gene discovery efforts may benefit from arraying probes for potentially all the expressed genes in the human genome. Even before the whole human genome is sequenced, the huge repertoire of expressed sequence tagged sites (ESTs) can be arrayed and studied.[48,53] Such a chip can be used to study gene expression during development, disease states and in normal metabolic homeostasis. Later, disease-targeted, user-adaptable arrays with probes for several hundred targets may be most useful in the clinical and biological research laboratories.

Other Anticipated Uses of DNA Microarrays

These include genetic linkage, forensic identification, pathogen identification, sequence analysis by hybridization of previously uncharacterized regions and comparative genome hybridization.[54–57] In addition, Single-Nucleotide Polymorphism (SNP)-based tests are adaptable to the microarray. The use of Southern blotting and detection of restriction fragment length polymorphism (RFLP) markers provided an initial relatively low resolution human genetic map for linkage analysis. These markers have largely been replaced by simple sequence repeat or microsatellite markers which are both simpler to run and are more highly polymorphic (more informative so fewer markers must be run to derive the same degree of confidence of linkage in gene discovery studies). The microsatellite map in human genetics is now quite dense, for example, Dibs[58] recently published the primer sequence and allele size ranges of 5264 microsatellite markers. Consequently, microsatellite genotyping has quickly become adopted in linkage analysis, disease association studies, and in the study of altered profiles diagnostic of certain cancers (evidence of loss of heterozygosity (LOH) or microsatellite instability (MSI). Thus, the microsatellite markers are the current method of choice for many molecular genetic research projects and provide the basis of a much more dense and powerful 'second generation' genetic map. However, microsatellite typing generally requires electrophoretic separation of the PCR-amplified markers in very high-resolution gels, a labor intensive and time-consuming process when slab-gel electrophoresis systems are used. Consequently, so-called 'third generation' marker sets are under

development at several genomic research institutes including the Whitehead Institute (Boston, MA) and GenSet (Paris, France).

The third-generation markers are single-nucleotide polymorphisms (SNP). SNP markers are fundamentally the same as the RFLP markers (base substitutions) but with improvements in detection methods, and are now identified and screened without the use of electrophoresis or Southern blotting. SNP markers can be detected in microplate or microarray-based formats. The inherent variability within the human genome results in base substitutions at approximately 1 kb intervals. When alternative alleles (base substitutions) have reached a frequency of 30% or greater, they can be very useful in genetic analysis. With the very high density of these markers in the genome, it is possible to determine regional haplotypes (ordered alleles that cosegregate along a common stretch of chromosome). Haplotype mapping helps compensate for the lower informativeness of each SNP marker relative to the microsatellite markers. Nevertheless, it is estimated that a 4-cM map of 750 SNP markers will be required to provide the same degree of genetic information as a 10-cM map of approximately 300 microsatellite markers.[56] The primary motivation for developing a third generation map is that the SNP markers are much more amenable to automation and detection on array-based systems. For example, the Whitehead Institute, Millennium and Affymetrix have an active program identifying and mapping SNP markers using genomic clones which are physically mapped. Similarly, GenSet in collaboration with Abbott Laboratories also have an SNP marker development program with the goal of identifying up to 30 000 SNP markers in the human genome by the end of 1999.

A set of SNP markers uniformly distributed across the genome would have additional applications besides linkage analysis. For example, the use of a 10 000 SNP marker set would permit a genome-wide scan for loss of heterozygosity (LOH) at a resolution of under 1 Mbase (average spacing of the SNP markers would be 300 000 nucleotides). This would be a higher resolution LOH scan than is currently possible using comparative genome hybridization (CGH) carried out using cytogenetic approaches.[57] By strategically positioning the SNP markers to deletion prone regions of the genome and/or to known locations of oncogenes, the utility of this approach to LOH scanning could be further improved. Tightly clustered SNP markers within large genes that lead to rare genetic disorders would provide an alternative method for a linkage-based strategy for disease diagnostics particularly in new mutation disorders (each family has a unique, uncharacterized mutation). For example, the dystrophin gene is approximately 2.4 Mbase in size and five microsatellite markers are used to follow intragenic recombination in families where either Duchenne or Becker muscular dystrophy are segregating and the responsible mutation fails to be detected using standard exon-screening tests.[59–61] Approximately 24 polymorphic SNP markers would be expected to be found within the dystrophin gene. These would naturally have tighter spacing than the current set of intragenic microsatellite markers. By scoring the alleles at the SNP markers, the haplotype of the disease-carrying allele could be identified. A linkage-based method of diagnosis may prove more reliable than a mutation screening-based method in the difficult case of accurately providing diagnosis when the disease-causing gene is very large. Similarly, complex disorders which have multiple genes involved might permit rapid diagnosis using a linkage-based approach. For example, it is estimated that there are 50 genes responsible for inherited deafness. By haplotyping in the vicinity of known genes that could contribute to this disorder, fewer family members might be required to identify which gene is involved for any one family.

Research and Development Issues in DNA Microarrays

Research and development issues in DNA microarrays that operate by hybridization have been the probes (the immobilized species), the targets (the solution-phase reverse complements of the surface probes), the surfaces, the detection labels, the hybridization conditions, and the detection principle. To our knowledge, 12 separate groups worldwide including our own to date have published the operating details of their microarrays. Probes are obtained by synthesis *in situ* or deposition of pre-synthesized molecules. Two main types of synthesis *in situ* have been described: light-based combinatorial[16,62] or physical combination.[29,63] Deposition of pre-synthesized oligonucleotides,[20,44,64,65] cDNAs[48–51] or nucleic acids with other base and backbone modifications such as PNAs[66–68] have been reported. Oligonucleotide probes appear to benefit from spacer arms that link the nucleic acids to the surface.[20,44,69] Targets have been single- or double-stranded DNA or RNA. A few investigators have explicitly examined varied hybridization conditions.[70,71] A variety of surfaces have been described, ranging from glass to silicon to plastic to paper to silanized

glass.[16,20,44,72,73] There is a multitude of detection labels, radioactive and light-based. Advantages accrue to the use of light-based detection labels. Detection labels have been based on fluorescence or chemiluminescence, and both primary (label on the target) and secondary (label on a molecule which binds a group on the target or duplex) detection are in use.[74] Detection devices include the proximal CCD,[64,75] CCD-microscope,[30,44] and PMT-confocal microscope[21–26] systems.

Discriminating single-base mutations from normals and especially heterozygotes is fraught with difficulties. We and others have seen unexpected and largely unexplainable hybridization patterns when a given target is hybridized to a series of probes with overlapping sequences.[44,76,77] Many examples have concentrated on the difficulties of detecting point mutations. However, similar statements could be made about present techniques for quantitating other important molecular genetic features, such as the absolute expression levels of particular mRNAs. In each case, hybridization alone, while powerful, is unlikely to be capable of providing a complete and unambiguous answer. We and others are turning to combined hybridization and reaction *in situ*: single-nucleotide extension (SNE),[43,78] PCR *in situ*,[79–81] reporter/quencher uncoupling chemistry (R/Q),[82] and ligase chain reaction (LCR)[83] among others. We designate microarrays that operate by hybridization that follows diffusion of targets to their cognate probes as first generation DNA microarrays. Second generation microarrays have one or more of the following features: active capture of targets, eg by electric field, and hybridization coupled to reaction. There has been a lot of empiricism, but also a considerable contribution to microarray design by fundamental studies.[35,44,45,69,72,73,76,84–87] The capture rate of a nucleic acid in solution, J (moles/cm^2), by its reverse complement immobilized on a solid surface, when the number of surface 'probe' molecules is in molar excess, is given by:[84,85]

$$J = (\pi\gamma D_3/4H)C_0t$$

Molecular biologists with experience in solution-phase hybridization studies understand the dependence on C_0t. This theoretical result places physical observables in the factors other than C_0t which determine capture and hybrid formation specifically at the solid-liquid interface. This J is the initial rate – eventually the capture saturates at an inverse exponential rate, but most detection can and does operate before that point is reached. The parameter γ incorporates all of the 3- and 2-dimensional diffusion effects, adsorption at the surface, probe density within a site, probe molecular length and target length. D_3 is the solution diffusion coefficient of the target. The smaller the target, the larger is D_3 and the more rapid its capture by the diffusive mechanism, supporting the empirical practice of fragmenting into smaller targets.[21–26] Fragmentation also minimizes secondary structure, an effect discussed later. H is the boundary layer, often in the length scale of tens of microns, over which the concentration changes from its bulk solution value to that right at the interface. Decreasing H, for example by stirring or by 'excluding volume' from the bulk solution with amphipathic organic molecules like PEG or phenol increases capture.[44,86,88] If a second mechanism of capture at the surface is included, like electric field, capture can be enhanced. This underlies another published approach, which is very rapid but employs more complex chip manufacture and operation.[70] The side-on orientation of DNA molecules on common silanized surfaces has been observed experimentally.[87] Further progress in rational design from biophysical and biochemical fundamental study is anticipated.

Non Microarray-based Parallel Mutation Detection Methods

DNA sequence variation detection methods share several conceptual issues, such as whether a known change is being detected or a region is being scanned for new mutations, whether the end user of the method is a clinical laboratory or a research laboratory, whether the change to be identified is a disease-causing mutation or a polymorphism, and whether the cellular sample is homozygous, hemizygous or heterozygous for the change.

A wide variety of PCR-based mutation detection methods have been described.[89,90] In general the mutation detection methods can be grouped according to whether they are scanning or allele specific. Within the scanning methods, they can be grouped according to which principle of operation they employ. Well-known applications for mutation screening include direct DNA sequence analysis, denaturing-gradient gel electrophoresis (DGGE),[91] single-stranded conformation polymorphism (SSCP)[92] and its variants dideoxyfingerprinting (ddF)[93] and restriction endonuclease fingerprinting,[94] heteroduplex analysis including chemical cleavage of mismatches (CCM)[95] and endonuclease

mismatch cleavage (EMC),[96] and RNase A cleavage.[97] Automated fluorescence-based DNA sequence analysis can generate 600 to 800 nucleotides of sequence data per reaction, and in some sense is parallel in that multiple gel electrophoresis lanes can be analyzed at the same time, as is true for SSCP, DGGE and some of the other methods. Allele-specific mutation/polymorphism detection methods include direct DNA sequence analysis,[98] reverse dot-blot with allele-specific oligonucleotide hybridization,[99] allele-specific PCR amplification,[100] oligonucleotide ligation amplification (OLA),[101] artificial introduction of restriction sites,[102] ligase chain reaction[103] and DNA minisequence analysis.[104–106] PCR-OLA has been applied to DNA diagnostics,[107] genetic mapping using biallelic markers[108] and YAC library screening.[109] A variation on PCR-OLA to increase sample throughput makes use of sequence-coded separation.[101] Fluorescence-based DNA minisequence analysis facilitates detection of mutated sequences.[110,111] Limited primer-extension techniques, primer-guided nucleotide incorporation, single-nucleotide primer extension (SNPE) or solid-phase minisequencing were developed previously to detect point mutations by single-base extension of the primer at the site of the mutation.[104–106] These procedures use a primer designed to hybridize just 5' of the nucleotide to be tested. The primer is extended by a single dye-labeled dideoxynucleotide, thereby indicating the identity of the target nucleotide in the template. Fluorescent extension products are then detected following electrophoresis on denaturing polyacrylamide gels. Fluorescence analysis of the incorporated dye tag reveals the identity of the template nucleotide immediately 3' to the primer site. There are recent examples of the application of the single nucleotide extension principle to the microarray format.[43,78]

Two newer methods are cleavase fragment length polymorphism (CFLP) and multiplex allele-specific diagnostic assay (MASDA). CFLP is a scanning methodology, while MASDA is an allele-specific methodology. CFLP is in the general category of secondary structure-specific mutation detection. Whilst SSCP reflects structure-specific mobilities, CFLP reflects structure-specific nuclease cleavage.[112] MASDA is a multistep process in which multiplex gene-specific PCR is followed by a forward dot-blot using a pool of radiolabeled ASO probes.[113] After the dot-blot, verification of which oligo in the pool hybridized to an array spot is performed by a biochemical fingerprinting assay.

To compare and contrast all of these methods with DNA microarray methods of mutation detection, it is clear that many non microarray methods are up and running now while DNA microarray methods have just passed proof-of-principle and are undergoing initial dissemination. (CE/CAE are discussed below.) Automated fluorescence-based DNA sequence analysis continues to have many desirable features. Gel-based methods that use multicolor fluorescence consolidate separation with detection and have the advantage that software analysis can be built in. The special conditions required for running assays such as DGGE and SSCP have made their implementation on machines designed originally for automated DNA sequencing more problematic. The absence of a cooling unit on these sequencers originally required SSCP gels to be run for long times at very low current to avoid denaturing the annealed species,[114] although other investigators have made modifications to their machines to provide refrigeration capabilities and permit shorter run times.[115,116] Some technical modifications were also necessary to reduce background problems in the implementation of the fluorescent versions of the CCM[117] and dideoxyfingerprinting assay.[118] Most recent versions of automated sequencers including capillary array electrophoresis equipment (CAE) have built-in temperature control, which permits SSCP and other temperature-dependent applications to be accomplished. Many of these assays can be done using fluorescently-labeled reactions and appropriate detection systems, resulting in automated data collection and semi-automated analysis.[119] In most cases, the use of fluorescence in these assays is limited only by the ability to label the species in an efficient manner without causing background problems.[120] How robust each of these methods are (accuracy, reproducibility and reliability) is of concern to end-users and has limited the acceptance of any one scanning or allele-specific method in molecular genetics laboratories. Microarrays have the inherent advantage that they can do both scanning and allele-specific mutation detection while consolidating the reaction step with detection, thus skipping a distinct separation step. Since physical localization is intrinsic to a spatially addressable array, there is no electrophoretic separation. Currently, microarrays avoid radioactivity and are readily operated with multicolor detection, which increases the amount of information per spot and per unit time. One major distinction between mutation detection methods in the future may be microarrays *vs* homogeneous

solution arrays that perform consolidated reaction and detection with no separation step (for example, 'molecular beacon'/Reporter-Quencher systems).[38,82,121] Since these homogeneous assays are still under development, how parallel, small and fast these are in comparison to microarrays remain to be determined. Mass spectrometry-based methods have also been proposed and are under development.[31–33]

Non Microarray-based Parallel Gene Expression Profiling

Analysis of mRNA, in particular differences in expression patterns between cell types, is an area of intense interest. At this time this is almost exclusively a research laboratory enterprise, not one for the clinical molecular diagnostics laboratory. Several common methods of RNA analysis are not inherently designed as parallel molecular genetic assays, such as northern blot analysis, nuclease protection assays and RT-PCR. Beyond the DNA microarray work described in the preceding section, there are a number of parallel gene expression profiling methods in use in many laboratories. These include

(1) subtractive hybridization and its variants, recently reviewed in Sagerstrom *et al*,[122] in which comparison with other established methods was discussed;

(2) differential display and its variants;

(3) differential screening of cDNA libraries;

(4) large-scale sequencing of cDNA clones or 3'-end fragments thereof;

(5) representational difference analysis (RDA); and

(6) serial analysis of gene expression (SAGE).

Subtractive hybridization (SH), differential display, differential screening, large-scale sequencing and RDA have been in the literature for several years, reviewed recently and are in use in many laboratories. Recent variations of these methods to improve their performance have been described. Improvements in subtractive hybridization may come from positive selection and so-called suppressive hybridization methods.[123] Elements of suppressive hybridization have also been combined recently with elements of differential screening. A variant of differential display PCR has been published recently called ordered dd-PCR.[124] Normalized or self-subtracted libraries appear to enhance performance in large scale DNA sequencing.[122] In commercial research labs, large-scale cDNA sequencing has been reported to identify many novel cDNAs. Much has been learned from the SAGE method.[125–127] SAGE works by cloning and sequencing strings of 9 to 11 bp fragments from cDNA libraries, and uses statistical sequence analysis first to assign the fragments uniquely to transcripts and then to calculate abundance. Many of the findings from classic hybridization studies in the 1970s concerning the number and abundance classes of mRNAs in cells have been confirmed;[122,128] 86% of transcripts by number are present at less than 5 copies per cell. Of the approximately 15 000 distinct transcripts in a cell, 1–2% differ in cells of closely related origins. If cells of more disparate origins are compared, the percentage of differentially expressed transcripts may rise to 3–5%. Differentially expressed transcripts occur in all abundance classes. SAGE is unfortunately not high throughput and thus is not amenable to the clinical laboratory.

Capillary Array Electrophoresis Methods for DNA Sequence Analysis and Microsatellite Genotype Analysis

Recently, fluorescence detection systems have been coupled with the speed and convenience of capillary electrophoresis (CE) systems which have potential for rapid, high-throughput genetic analysis without the constraints and disadvantages of slab-gel based systems. These systems have already been used to analyze fluorescent sequencing reactions and a variety of dye-labeled fragments.[129,130] Multiple arrays of capillaries and advances in separation matrices will further enhance resolution, sample throughput and decrease run times. Capillary technology also has the potential to be engineered on a microchip scale, decreasing electrophoresis times by an order of magnitude.[131–133]

Mathies and co-workers developed a laser-excited, confocal fluorescence imaging system for DNA sequencing or mapping fragments in a planar array of several capillaries simultaneously, a technology known as capillary array electrophoresis (CAE).[134,135] CAE offers an excellent opportunity to develop instrumentation capable of striking increases in efficiency and convenience for DNA analysis. Molecular Dynamics, extending the work in CAE, is developing instrument

systems capable of scanning across 48–96 parallel capillaries.[136,137] This could increase throughput of fluorescence DNA sequencing and mapping by over an order of magnitude.

Labor-intensive steps of fragment separations are eliminated by CAE. Slab gel preparation is replaced by automated capillary filling using a low-viscosity gel matrix, refillable capillaries and a pressurized gel-filling station.[138] The process is complete within 10 minutes without user intervention. Gel loading, another tedious, labor-intensive step, is replaced with electrokinetic injection of 48–96 samples at a time. Capillaries are then purged and refilled using the gel-filling station on the instrument. Thus, the increase in sample throughout is not limited to electrophoresis time, but all steps of gel preparation and sample loading. Use of 4-color chemistries has been used in DNA sequencing and permits detection of multiple genetic markers with overlapping size ranges. Multiplexing strategies increase sample throughput and conserve reagent costs when done during the PCR amplification step.

Estimated time from capillary filling, sample injection, separation and refilling of the capillary for the next run is under 1.5 hours.[138,139] With continuous operation (16 runs per day), 1536 samples could be processed per day by CAE. By using the 4-color multiplex strategy that has been applied to slab-gel electrophoresis systems, 27 648 genotypes could be determined per day (1536 samples $\times$ 3 colors of labels $\times$ 6 loci/color) compared to a maximum of 4 runs in 24 hours on current slab-gel systems. The CAE approach translates into 2304 genotype analyses per day (128 samples $\times$ 3 colors of labels $\times$ 6 loci/color). Thus, CAE systems would have a ten-fold greater throughput and require much less hands-on time for operation.

Conclusions

We have witnessed the initial impact of parallel molecular genetic analysis, increasing sample throughput and facilitating semi-automated data analysis. More importantly, these new technologies, such as DNA microarrays and capillary array electrophoresis, replace standard molecular biology platforms and may lead to further technological advances in genetic analysis. In addition, recent developments in scanner technology promise to reduce the time needed to obtain large amounts of information at the same time.

Some specific comparative strengths and weaknesses of analytic technologies for mutation detection and gene expression profiling have been listed in the individual sections. Here we summarize a framework for decision making in selecting available technology options, which is tabulated in Table 1 for seven of the most common molecular genetic tasks. Our philosophy is to list the technology options for which there is a body of literature and to identify the criteria for

Table 1 Parallel molecular genetic analysis – framework for decision making in selecting available technology options

Molecular genetic task[1]	*Analytic technology options*	*Criteria for selection by end users*[3]
A Mutation detection, allele-specific	1, 2, 3, 4, 5, 6, 7	A, B, C, D, E, F
B Mutation detection, scanning known gene(s)	1, 3, 4	A, D, E, F
C Sequence analysis, previously unknown region	1, 3, 4	A, D
D Polymorphism scoring, eg for linkage		
microsatellite (MS)	3, 4, 5, 7	A, B, C
single nucleotide polymorphism (SNP)	1, 2	A, C, D
E Gene expression profiling	1	A, C, D, E, F
F Gene copy number determination	1, 4, 5	A, C

Key:

A 18, 19, 31–33, 35, 36, 43, 45, 63, 78, 117	1 DNA microarray, 1st generation	A Cost – upfront
B 17, 21–29, 31, 34, 36, 37, 39, 42, 55, 64–66, 70–72, 75	2. DNA microarray, 2nd generation	B Cost – ongoing operation
C 37, 39, 129, 133, 137	3 Rapid gel-based systems	C How parallel
D (MS) 31, 131, 132, 138, 139	4 Capillary array electrophoresis	D How high-throughput
D (SNP) 20, 30	5 Homogeneous solution assay	E How adaptable
E 46–54	6 Mass spectroscopy	F How generalizable
F 57, 83, 121	7 Hybrid of the above	

[1]Tasks range from allele-specific to genome-wide; additional uses in research and development include forensic identification, pathogen identification, microsatellite instability analysis and DNA methylation analysis.
[2]References based primarily on the peer-reviewed literature up to 01/31/1998.
[3]End users include biomedical research laboratories, molecular diagnostic clinical laboratories and the pharmaceutical industry.

selection by end users. The literature in this area is exploding rapidly, and our tabulation is as of 30 January 1998. Updates will clearly be needed in the future, and other recent reviews offer alternative opinions.[52,140–143] The technology option which is best will be driven by different criteria for different users. The relative importance of upfront and ongoing operational costs and how adaptable and generalizable a technology is will depend on whether one is in an academic biomedical research center, a clinical molecular diagnostics laboratory or the pharmaceutical industry, among other venues. By generalizable we mean able to be used by any laboratory with current molecular genetics expertise. By adaptable we mean able to be used for new genes or by multiple users with different interests. In first generation DNA microarray technology, the upfront costs for an arrayer and a scanner system are of the order of $200 000–250 000. Chip costs vary between premade and marketed and making one's own. High volume use keeps the per chip cost low, and then operator time and sample preparation become significant ongoing operational costs. Capillary array electrophoresis units are expected to cost under $200 000 initially, and operating costs will be comparable to gel-based systems. In summary, in Table 1 we provide a framework by which readers can make their own best decisions.

We anticipate that in the future consolidation of function and miniaturization will continue. Consolidation of the functions of sample preparation, target amplification and labelling, real time detection, analysis and informatics will result in very powerful tools. As the Human Genome Initiative meets its objectives, we may find a return to a focus on parallel molecular genetic technologies for specific diseases and conditions, eg all cardiovascular disease susceptibility genetic changes, all molecular alterations present in leukemia, all genetic variations relevant to drug metabolism, etc.

Acknowledgements

This work was supported in part by a grant from the Ethel Brown Foerderer Fund for Excellence, by NIH grants P60-HL38632, P30-HG00425, R01-DK16691, and SBIR R43NS/MH34589, a Special Opportunity Award of the Whitaker Foundation and by the Nemours Foundation.

References

1 Lander ES: The new genomics: global views of biology. *Science* 1996; **274**: 536–539.
2 Cotton RG: Slowly but surely towards better scanning for mutations. *Trends Genet* 1997; **13**: 43–46.
3 Fodor SPA: Massively parallel genomics. *Science* 1997; **277**: 393–395.
4 Goffeau A: DNA technology: molecular fish on chips. *Nature* 1997; **385**: 202–203.
5 Cheng J, Fortina P, Surrey S, Kricka LJ, Wilding P: Microchip-based devices for molecular diagnosis of genetic diseases. *Mol Diagnosis* 1996; **1**: 1–18.
6 Strachan T, Abitbol M, Davidson D, Beckmann JS: A new dimension for the human genome project: towards comprehensive expression maps. *Nat Genet* 1977; **16**: 126–132.
7 Anderson NG, Anderson NL: Twenty years of 2D electrophoresis: past, present and future. *Electrophoresis* 1996; **17**: 443–453.
8 Roepstorff P: Mass spectrometry in protein studies from genome to function. *Curr Opin Biotech* 1997; **8**: 6–13.
9 Shevchenko A, Jensen ON, Podtelejnikov AV *et al*: Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. *Proc Natl Acad Sci USA* 1996; **93**: 14440–14445.
10 Northrup MA, Ching MT, White RM, Watson RT: DNA amplification in a microfabricated reaction chamber. In *Transducers '93, 7th International Conference on Solid State Sensor Actuators*. Yokohama, Japan, June 7–10, 1993, pp 924–926.
11 Kricka LJ, Nozaki O, Heyner S, Garside WT, Wilding P: Application of a microfabricated device for evaluating sperm function. *Clin Chem* 1993; **39**: 1944–1947.
12 Wilding P, Shoffner MA, Kricka LJ: PCR in a silicon microstructure. *Clin Chem* 1994; **40**: 1815–1818.
13 Benton D: Bioinformatics: principles and potential of a new multidisciplinary tool. *Trends Biotech* 1996; **14**: 261–272.
14 Palsson BO: What lies beyond bioinformatics? *Nat Biotechnol* 1997; **15**: 3–4.
15 Rawlings CJ, Searls DB: Computational gene discovery and human disease. *Curr Opin Genet Develop* 1997; **7**: 416–423.
16 Fodor SPA, Read JL, Pirrung MC, Stryer L, Lu AT, Solas D: Light-directed, spatially addressable parallel chemical synthesis. *Science* 1991; **251**: 767–773.
17 Southern EM, Maskos U, Elder JK: Analyzing and comparing nucleic acid sequences by hybridization to arrays of oligonucleotides: evaluation using experimental models. *Genomics* 1992; **13**: 1008–1017.
18 Maskos U, Southern EM: A novel method for the analysis of multiple sequence variants by hybridization to oligonucleotides. *Nucleic Acids Res* 1993; **21**: 2267–2268.
19 Maskos U, Southern EM: A novel method for the parallel analysis of multiple mutations in multiple samples. *Nucleic Acids Res* 1993; **21**: 2269–2270.
20 Guo Z, Guilfoyle RA, Thiel AJ, Wang R, Smith LM: Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports. *Nucleic Acids Res* 1994; **22**: 5456–5465.
21 Pease AC, Solas D, Sullivan EJ, Cronin MT, Holmes CP, Fodor SPA: Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc Natl Acad Sci USA* 1994; **91**: 5022–5026.

22 Lipshutz RJ, Morris D, Chee M *et al*: Using oligonucleotide probe arrays to access genetic diversity. *Biotechniques* 1995; **19**: 442–447.
23 Cronin MT, Fucini SMK, Masino RS, Wespi RM, Miyada CG: Cystic fibrosis mutation detection by hybridization to light-generated DNA probe arrays. *Hum Mut* 1996; **7**: 244–255.
24 Chee M, Yang R, Hubbell E *et al*: Accessing genetic information with high-density DNA arrays. *Science* 1996; **274**: 610–614.
25 Kozal MJ, Shah N, Shen N *et al*: Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. *Nat Med* 1996; **2**: 753–759.
26 Hacia JG, Brody LC, Chee MS, Fodor SPA, Collins FS: Detection of heterozygous mutations in BRCA1 using high density oligonucleotide arrays, two colour fluorescence analysis. *Nat Genet* 1996; **14**: 441–447.
27 Yershov G, Barsky V, Belgovskiy A *et al*: DNA analysis and diagnostics on oligonucleotide microchips. *Proc Natl Acad Sci USA* 1996; **93**: 4913–4918.
28 Drobyshev A, Moligina N, Shik V, Pobedimskaya D, Yershov G, Mirzabekov A: Sequence analysis by hybridization with oligonucleotide microchip: identification of β-thalassemia mutations. *Gene* 1997; **188**: 45–52.
29 Southern EM: DNA chips: analysing sequence by hybridization to oligonucleotides on a large scale. *Trends Genet* 1996; **12**: 110–115.
30 Guo A, Liu Q, Smith LM: Enhanced discrimination of single nucleotide polymorphisms by artificial mismatch hybridization. *Nat Biotech* 1997; **15**: 331–335.
31 Little DP, Braun A, O'Donnell MJ, Koster H: Mass spectrometry from miniaturized arrays for full comparative DNA analysis. *Nat Med* 1997; **3**: 1413–1416.
32 Higgins GS, Little DP, Koster H: Competitive oligonucleotide single-base extension combined with mass spectrometric detection for mutation screening. *Biotechniques* 1997; **23**: 710–714.
33 Griffin TJ, Tang W, Smith LM: Genetic analysis by peptide nucleic acid affinity MALDI-TOF mass spectrometry. *Nat Biotech* 1997; **15**: 1368–1372.
34 Head SR, Rogers Y-H, Parikh K *et al*: Nested genetic bit analysis (N-GBA) for mutation detection in the p53 tumor suppressor gene. *Nucleic Acids Res* 1997; **25**: 5065–5071.
35 Edman CF, Raymond DE, Wu DJ *et al*: Electric filed directed nucleic acid hybridization on microchips. *Nucleic Acids Res* 1997; **25**: 4907–4914.
36 Hacia JG, Makalowski W, Edgemon K *et al*: Evolutionary sequence comparisons using high-density oligonucleotide arrays. *Nat Genet* 1998; **18**: 155–158.
37 Stein A, Hill SA, Cheng Z, Bina M: Simple miniaturized gel system for DNA sequence analysis. *Nucleic Acids Res* 1998; **26**: 452–455.
38 Tyagi S, Bratu DP, Kramer FR: Multicolor molecular beacons for allele discrimination. *Nat Biotech* 1998; **16**: 49–53.
39 Drmanac S, Kita D, Labat I *et al*: Accurate sequencing by hybridization for DNA diagnostics and individual genomics. *Nat Biotech* 1998; **16**: 54–58.
40 Shoffner MA, Cheng J, Hvichia GE, Kricka LJ, Wilding P: Chip PCR (I): surface passivation of micro-fabricated silicon-glass chips for PCR. *Nucleic Acids Res* 1996; **24**: 375–379.
41 Cheng J, Shoffner MA, Hvichia GE, Kricka LJ, Wilding P: Chip PCR (II): investigation of different PCR amplification systems in micro-fabricated silicon-glass chips. *Nucleic Acids Res* 1996: **24**: 380–385.
42 Taylor TB, Winn-Deen ES, Picozza E, Woudenberg TM, Albin M: Optimization of the performance of the polymerase chain reaction in silicon-based microstructure. *Nucleic Acids Res* 1997; **25**: 3164–3168.
43 Pastinen T, Kurg A, Metspalu A, Peltonen L, Syvanen A-C: Minisequencing: a specific tool for DNA analysis and diagnostics on oligonucleotide arrays. *Genome Res* 1997; **7**: 606–614.
44 Sanguedolce LA, Chan V, McKenzie S, Surrey S, Fortina P, Graves D: Fundamental studies of DNA adsorption and hybridization on solid surfaces. In: Dinh S, De Nuzzio J (eds). *Intelligent Materials and Novel Concepts for Controlled-release Technologies*. American Chemical Society Symposium Series. 1998, in press. Academic Press.
45 Sanguedolce LA: Fundamental studies of DNA attachment and hybridization at a solid-liquid interface. University of Pennsylvania, 1997, PhD thesis.
46 Lockhart DJ, Dong H, Byrne MC *et al*: Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotech* 1996; **14**: 1675–1680.
47 Wodicka L, Dong H, Mittman M, Ho M-H, Lockhart DJ: Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat Biotech* 1997; **15**: 1359–1367.
48 DeRisi J, Penland L, Brown PO *et al*: Use of a cDNA microarray to analyze gene expression patterns in human cancer. *Nat Genet* 1996; **14**: 457–460.
49 Schena M, Shalon D, Davis RW, Brown PO: Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995; **270**: 467–470.
50 Schena M, Shalon D, Heller R, Chai A, Brown PO, Davis RW: Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc Natl Acad Sci USA* 1996; **93**: 10614–10619.
51 Shalon D, Smith SJ, Brown PO: A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res* 1996; **6**: 639–645.
52 Castellino AM: When the chips are down. *Genome Res* 1997; **7**: 943–946.
53 Ferguson JA, Boles TC, Adams CP, Walt DR: Fiberoptic DNA biosensor microarray for the analysis of gene-expression. *Nat Biotech* 1996; **14**: 1681–1684.
54 Heller RA, Schena M, Chai A *et al*: Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. *Proc Natl Acad Sci USA* 1997; **94**: 2150–2155.
55 Drmanac R, Drmanac S, Strezoska Z *et al*: DNA sequence determination by hybridization: a strategy for efficient large-scale sequencing. *Science* 1993; **260**: 1649–1652.
56 Kruglyak L: The use of a genetic map of biallelic markers in linkage studies. *Nat Genet* 1997; **17**: 21–24.

57 Solinas-Toldo S, Lampel S, Stilgenbauer S *et al*: Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosom Cancer* 1997; **20**: 399–407.
58 Dibs C, Faure S, Fizames C *et al*: A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 1996; **380**: 152–154.
59 Chamberlain JS, Gibbs RA, Rainier JE, Nguyen PN, Caskey CT: Deletions screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. *Nucleic Acids Res* 1988; **16**: 11141–11156.
60 Beggs AH, Koerig M, Boyce FM, Kurkel LML: Detection of 98% of DMD/BMD gene deletions by polymerase chain reaction. *Hum Genet* 1990; **86**: 45–48.
61 Clements PR, Fenwick RG, Chamberlain JS *et al*: Carrier detection and prenatal diagnosis in Duchenne and Becker muscular dystrophy families using dinucleotide repeat polymorphisms. *Am J Hum Genet* 1991; **49**: 951–960.
62 McGall G, Labadie J, Brock P, Wallraff G, Nguyen T, Hinsberg W: Light-directed synthesis of high density oligonucleotide arrays using semiconductor photoresists. *Proc Natl Acad Sci USA* 1996; **93**: 13555–13560.
63 Nilsson M, Malmgren H, Samiotaki M, Kwiatkowski M, Chowdhary BP, Landegren U: Padlock probes: circularizing oligonucleotides for localized DNA detection. *Science* 1994; **265**: 2085–2088.
64 Eggers MD, Hogan ME, Reich RK *et al*: Gensensors: microfabricated devices for automated DNA sequence analysis. In: *Advances in DNA Sequencing Technology*. Proceedings SPIE 1993; pp 113–126.
65 Beattie K, Eggers M, Shumaker J *et al*: Genosensor technology. *Clin Chem* 1993; **39**: 719–722.
66 Weiler J, Gausepohl H, Hauser N, Jensen O, Hoheisel JD: Hybridisation based DNA screening on peptide nucleic acid (PNA) oligomer arrays. *Nucleic Acids Res* 1997; **25**: 2792–2799.
67 Wittung P, Nielsen PE, Buchardt O, Egholm M, Norden B: DNA-like double helix formed by peptide nucleic acid. *Nature* 1994; **368**: 561–563.
68 Nguyen H-K, Auffray P, Asseline U, Dupret D, Thuong NT: Modification of DNA duplexes to smooth their thermal stability independently of their base content for DNA sequencing by hybridization. *Nucleic Acids Res* 1997; **25**: 3059–3065.
69 Maskos U, Southern EM: Oligonucleotide hybridisations on glass supports: a novel linker for oligonucleotide synthesis, hybridisation properties of oligonucleotides synthesized *in situ*. *Nucleic Acids Res* 1992; **20**: 1679–1684.
70 Sosnowski RG, Tu E, Butler WF, O'Connell JP, Heller MJ: Rapid determination of single base mismatch mutations in DNA hybrids by direct electric field control. *Proc Natl Acad Sci USA* 1997; **94**: 1119–1123.
71 Dubiley S, Kirillov E, Lysov Y, Mirzabekov A: Fractionation, phosphorylation and ligation on oligonucleotide microchips to enhance sequencing by hybridization. *Nucleic Acids Res* 1997; **25**: 2259–2265.
72 Matson RS, Rampal J, Pentoney J, Stephen L, Anderson PD, Coassin P: Biopolymer synthesis on polypropylene supports: oligonucleotide arrays. *Anal Biochem* 1995; **224**: 110–116.
73 Chrisey LA, Lee GU, O'Ferrall CE: Covalent attachment of synthetic DNA to self-assembled monolayer films. *Nucleic Acids Res* 1996; **24**: 3031–3039.
74 Mansfield ES, Worley JM, McKenzie SE, Surrey S, Rappaport EF, Fortina P: Nucleic acid detection using non-radioactive labeling methods. *Mol Cell Probes* 1995; **9**: 145–156.
75 Lamture JB, Beattie KL, Burke BE *et al*: Direct detection of nucleic acid hybridization on the surface of a charge coupled device. *Nucleic Acids Res* 1994; **22**: 2121–2125.
76 Shchepinov MS, Case-Green SC, Southern EM: Steric factors influencing hybridization of nucleic acids to oligonucleotide arrays. *Nucleic Acids Res* 1997; **25**: 1155–1161.
77 Lima WF, Monia BP, Ecker DJ, Freier SM: Implication of RNA structure on antisense oligonucleotide hybridization kinetics. *Biochemistry* 1992; **31**: 12055–12061.
78 Shumaker JM, Metspalu A, Caskey CT: Mutation detection by solid phase primer extension. *Hum Mut* 1996; **7**: 346–354.
79 Bryant N: Use of PCR method expands to include broad range of diverse applications. *Genetic Engineering News* 1997; **1**: 9,31.
80 Adams CP, Kron SJ: Method for performing amplification of nucleic acid with two primers bound to a single solid support. Mosaic Technologies Inc, Whitehead Institute for Biomedical Research, 1997, USA patent 5641658.
81 Lockley AK, Jones CG, Bruce JS, Franklin SJ, Bardsley RG: Colorimetric detection of immobilised PCR products generated on a solid support. *Nucleic Acids Res* 1997; **25**: 1313–1314.
82 Tyagi S, Kramer FR: Molecular beacons – probes that fluoresce upon hybridization. *Nat Biotech* 1996; **14**: 303–308.
83 Cheng J, Shoffner MA, Mitchelson KR, Kricka LJ, Wilding P: The examination of ligase chain reaction (LCR) products amplified in the silicon chip using entangled solution capillary electrophoresis (ESCE). *J Chromatogr A* 1996; **732**: 151–158.
84 Chan V, Graves DJ, McKenzie SE: The biophysics of DNA hybridization with immobilized oligonucleotide probes. *Biophys J* 1995; **69**: 2243–2255.
85 Chan V, Graves DJ, Fortina P, McKenzie SE: Adsorption and surface diffusion of DNA oligonucleotides at liquid/solid interfaces. *Langmuir* 1997; **13**: 320–329.
86 Chan V: Reduction of dimensionality processes of DNA hybridization with immobilized oligonucleotide probes. University of Pennsylvania, 1997, PhD thesis.
87 Chan V, McKenzie SE, Surrey S, Fortina P, Graves DJ: The effect of hydrophobicity and electrostatics on adsorption and surface diffusion of DNA oligonucleotides at liquid/solid interfaces. *J Colloid Interface Sci* 1998; **203**: 197–207.
88 Pontius BW, Berg P: Rapid renaturation of complementary DNA strands mediated by cationic detergents: a role for high probability binding domains in enhancing the kinetics of molecular assembly processes. *Proc Natl Acad Sci USA* 1991; **88**: 8237–8241.
89 Grompe M: The rapid detection of unknown mutations in nucleic acids. *Nat Genet* 1993; **5**: 111–117.

90 Chehab FF: Molecular diagnostics: past, present and future. *Hum Mut* 1993; **2**: 331–337.
91 Sheffield VC, Cox DR, Lerman LS, Myers RM: Attachment of a 40-base-pair G + C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. *Proc Natl Acad Sci USA* 1989; **86**: 232–236.
92 Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T: Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc Natl Acad Sci USA* 1989; **86**: 2766–2770.
93 Sarkar G, Yoon H-S, Sommer SS: Dideoxy fingerprinting (ddF): a rapid and efficient screen for the presence of mutations. *Genomics* 1992; **13**: 441–443.
94 Liu Q, Sommer SS: Restriction endonuclease fingerprinting: a sensitive method for screening mutations in long, contiguous segments of DNA. *Biotechniques* 1995; **18**: 470–477.
95 Cotton RGH, Rodrigues NR, Campbell RD: Reactivity of cytosine and thymine in single-base pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. *Proc Natl Acad Sci USA* 1988; **85**: 4397–4401.
96 Youil R, Kemper BW, Cotton RG: Screening for mutations by enzyme mismatch cleavage with T4 endonuclease. *Proc Natl Acad Sci USA* 1995; **92**: 87–91.
97 Myers RM, Larin Z, Maniatis T: Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA duplexes. *Science* 1985; **230**: 1242–1246.
98 McBride LJ, Koepf SM, Gibbs RA *et al*: Automated DNA sequencing methods involving polymerase chain reaction. *Clin Chem* 1989; **5**: 2196–2201.
99 Saiki RK, Walsh PS, Levenson CH, Erlich HA: Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. *Proc Natl Acad Sci USA* 1989; **86**: 6230–6234.
100 Newton CR, Graham A, Heptinstall LE *et al*: Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res* 1989; **17**: 2503–2516.
101 Grossman PD, Bloch W, Brinson E *et al*: High-density multiplex detection of nucleic acid sequences: oligonucleotide ligation assay and sequence-coded separation. *Nucleic Acids Res* 1994; **22**: 4527–4534.
102 Gasparini P, Bonizzato A, Dognini M, Pignatti PF: Restriction site generating-polymerase chain reaction (RG-PCR) for the probeless detection of hidden genetic variation: application to the study of some common cystic fibrosis mutations. *Mol Cell Probes* 1992; **6**: 1–7.
103 Wu DY, Wallace DB: The ligation amplification reaction (LAR) – Amplification of specific DNA sequences using sequential rounds of template-dependent ligation. *Genomics* 1989; **4**: 560–569.
104 Syvanen A-C, Aalto-Setala K, Harju L, Kontula K, Soderlund H: A primer-guided nucleotide incorporation assay in the genotyping of apolipoprotein E. *Genomics* 1990; **8**: 684–692.
105 Kuppuswamy MN, Hoffmann JW, Kasper CK, Spitzer SG, Groce SL, Bajaj SP: Single nucleotide primer extension to detect genetic diseases: Experimental application to hemophilia B (factor IX), cystic fibrosis genes. *Proc Natl Acad Sci USA* 1991; **88**: 1143–1147.
106 Jalanko A, Kere J, Savilahti E *et al*: Screening for defined cystic fibrosis mutations by solid-phase minisequencing. *Clin Chem* 1992; **38**: 39–43.
107 Nickerson DA, Kaiser R, Lappin S, Stewart J, Hood L, Landegren U: Automated DNA diagnostics using an ELISA-based oligonucleotide ligation assay. *Proc Natl Acad Sci USA* 1990; **87**: 8923–8927.
108 Nickerson DA, Whitehurst C, Boysen C, Charmley P, Kaiser R, Hood L: Identification of cluster of biallelic polymorphic sequence-tagged sites (pSTSs) that generate highly informative and automatable markers for genetic linkage mapping. *Genomics* 1992; **12**: 377–387.
109 Kwok P-Y, Gremaud MF, Nickerson DA, Hood L, Olson MV: Automatable screening of yeast artificial-chromosome libraries based on the oligonucleotide-ligation assay. *Genomics* 1992; **12**: 935–942.
110 Kobayashi M, Rappaport E, Blasband A *et al*: Fluorescence-based DNA minisequence analysis for detection of single-base changes in genomic DNA. *Mol Cell Probes* 1995; **9**: 175–182.
111 Fahy E, Nazarbaghi R, Zomorrodi M *et al*: Multiplex fluorescence-based primer extension method for quantitative mutation analysis of mitochondrial DNA and its diagnostic application for Alzheimer's disease. *Nucleic Acids Res* 1997; **25**: 3102–3109.
112 Brow MA, Oldenburg MC, Lyamichev V *et al*: Differentiation of bacterial 16S rRNA genes and intergenic regions and Mycobacterium tuberculosis katG genes by structure-specific endonuclease cleavage. *J Clin Microb* 1996; **34**: 3129–3137.
113 Shuber AP, Michalowsky LA, Nass GS *et al*: High throughput parallel analysis of hundreds of patient samples for more than 100 mutations in multiple disease genes. *Hum Mol Genet* 1997; **6**: 337–347.
114 Ellison J, Dean M, Goldman D: Efficacy of fluorescence-based PCR-SSCP for detection of point mutations. *Biotechniques* 1993; **15**: 684–691.
115 Iwahana H, Yoshimoto K, Mizusawa N, Kudo E, Itakura M: Multiple fluorescence-based PCR-SSCP analysis. *Biotechniques* 1994; **16**: 296–305.
116 Iwahana H, Adzuma K, Takahashi Y, Katashima R, Yoshimoto K, Itakura M: Multiple fluorescence-based PCR-SSCP analysis with postlabeling. *PCR Meth Appl* 1995; **4**: 275–282.
117 Haris II, Green PM, Bentley DR, Giannelli F: Mutation detection by fluorescent chemical cleavage: application to hemophilia B. *PCR Meth Appl* 1994; **3**: 268–271.
118 Ellison J, Squires G, Crutchfield C, Goldman D: Detection of mutations and polymorphisms using fluorescence-based dideoxy fingerprinting F-ddF. *Biotechniques* 1994; **17**: 742–753.
119 Fortina P, Conant R, Parrella T *et al*: Fluorescence-based, multiplex allele-specific PCR (MASPCR) detection of the ΔF508 deletion in the cystic fibrosis transmembrane conductance regulator (CFTR) gene. *Mol Cell Probes* 1992; **6**: 353–356.
120 Livak KJ, Flood SJ, Marmaro J, Giusti W, Deetz K: Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR Meth Appl* 1995; **4**: 357–362.

121 Wittwer CT, Herrmann MG, Moss AA, Rasmussen RP: Continuous fluorescence monitoring of rapid cycle DNA amplification. *Biotechniques* 1997; **22**: 130–138.
122 Sagerstrom CG, Sun BI, Sive HL: Subtractive cloning: past, present, and future. *Ann Rev Biochem* 1997; **66**: 751–783.
123 von Stein OD, Thies WG, Hofmann M: A high throughput screening for rarely transcribed differentially expressed genes. *Nucleic Acids Res* 1997; **25**: 2598–2602.
124 Matz M, Usman N, Shagin D, Bogdanova E, Lukyanov S: Ordered differential display: a simple method for systematic comparison of gene expression profiles. *Nucleic Acids Res* 1997; **25**: 2541–2542.
125 Zhang L, Zhou W, Velculescu VE *et al*: Gene expression profiles in normal and cancer cells. *Science* 1997; **276**: 1268–1272.
126 Velculescu VE, Zhang L, Vogelstein B, Kinzler KW: Serial analysis of gene expression. *Science* 1995; **270**: 484–487.
127 Velculescu VE, Zhang L, Zhou W *et al*: Characterization of the yeast transcriptome. *Cell* 1997; **88**: 243–251.
128 Wetmur JG: DNA probes: applications of the principles of nucleic acid hybridization. *Crit Rev Biochem Mol Biol* 1991; **26**: 227–259.
129 Luckey J, Drossman H, Kostichka T, Smith LM: High-speed DNA sequencing by capillary gel electrophoresis. *Meth Enzymol* 1993; **218**: 154–172.
130 Liu MS, Rampal S, Evangelista RA, Lee GC, Chen FT: Detection of amplified Y chromosome-specific sequence by capillary electrophoresis with laser-induced fluorescence. *Fertil Steril* 1995; **64**: 447–451.
131 Woolley AT, Mathies RA: Ultra-high-speed DNA fragment separations using microfabricated capillary array electrophoresis chips. *Proc Natl Acad Sci USA* 1994; **91**: 11348–11352.
132 Woolley A, Mathies RA: Ultra-high-speed DNA sequencing using capillary electrophoresis chips. *Anal Chem* 1995; **67**: 3676–3680.
133 Fung EN, Young ES: High-speed DNA sequencing by using mixed poly(ethylene oxide) solutions in uncoated capillary columns. *Anal Chem* 1995; **67**: 1913–1919.
134 Mathies RA, Huang XC: Capillary array electrophoresis: an approach to high-speed, high-throughput DNA sequencing. *Nature* 1992; **359**: 167–169.
135 Clark SM, Mathies RA: High-speed parallel separation of DNA restriction fragments using capillary array electrophoresis. *Anal Biochem* 1993; **215**: 163–170.
136 Bashkin J, Marsh M, Barker D, Johnston R: DNA sequencing by capillary electrophoresis with a hydroxyethylcellulose sieving buffer. *Appl Theor Electrophor* 1996; **6**: 23–28.
137 Bashkin J, Solomon N, Erndt N *et al*: DNA sequencing with a multi-capillary array system. *Microbiology Comparative Genomics* 1997; **2**: A178.
138 Mansfield ES, Vainer M, Enad S *et al*: Sensitivity, reproducibility and accuracy in short tandem repeat genotyping using capillary array electrophoresis. *Genome Res* 1996; **6**: 893–903.
139 Mansfield ES, Vainer M, Harris DW *et al*: Rapid sizing of polymorphic microsatellite markers by capillary array electrophoresis. *J Chromatogr A* 1997; **781**: 295–305.
140 Ramsay G: DNA chips: State-of-the-art. *Nat Biotech* 1998; **16**: 40–44.
141 Schafer AJ, Hawkins JR: DNA variation and the future of human genetics. *Nat Biotech* 1998; **16**: 33–39.
142 Marshall A, Hodgson J: DNA chips. An array of possibilities. *Nat Biotech* 1998; **16**: 27–31.
143 Brown PO, Hartwell L: Genomics and human disease – variations on variation. *Nat Genet* 1998; **18**: 91–93.